ELSEVIER

Journal of Chromatography A, 706 (1995) 31–36

JOURNAL OF CHROMATOGRAPHY A

"Chromatogram generator" chromatogram modelling software

A.V. Pirogov*, O.N. Obrezkov, O.A. Shpigun

Chemistry Department, M.V. Lomonosov Moscow State University, Lenin Hills, GSP-3, 119899 Moscow, Russian Federation

Abstract

Software called "Chromatogram Generator" was developed for IBM-compatible computers. This software allows the approximation of chromatographic peaks with several mathematical functions and their superposition, the construction of series of chromatograms with various parameters (or with smooth variations of single or several parameters), the construction of the graphical image of a chromatogram with minimum size and the ability to be added to a chromatographic database.

1. Introduction

Recently, much attention has been paid to the development of computerized expert systems, which aid analysts in the choice of the analytical method for a particular application and the search for the optimum conditions. For this purpose, the computer should have an option of quantitative estimation of the quality of chromatograms and effect their comparison. This can be performed using various optimization criteria. A number of papers have discussed these criteria [1–5]. However, the estimation of the quality of a chromatogram made by a specialist in the field of chromatography still appears to be the most reasonable method. Commercially available expert systems for chromatographic analysis pay most attention to the data acquisition and handling and to the choice of the optimum analytical conditions. These systems do not have the option of chromatogram modelling for expert estimation of their quality. At present, widely used chromatographic expert system software such as DryLab (LC Resources) does not provide for these important options. The best way to solve the task of peak estimation is to build a system that is able to learn about the shape of the chromatogram according to human experience.

To evaluate such chromatographic modelling software, a lot of chromatograms should be available, and the number of peaks, their shape, resolution, analysis time and a number of other parameters should vary over wide ranges to provide representative and reliable estimates. One of the ways to solve this problem is to create special software that can generate the chromatographic peaks, use them to build chromatograms and treat the resulting information.

2. Software specifications and hardware requirements

The software developed is called "Chromatogram Generator". It is written in C++ programming language (Borland International, Scotts Valley, CA, USA). The following external libraries were used: Borland Graphics Interface

* Corresponding author.

0021-9673/95/$09.50
SSDI 0021-9673(94)01243-1

(BGI) graphical library (EGAVGA.BGI) and IMPLODE compression library (PkWare, version 1.02). "Chromatogram Generator" requires an IBM PC/AT or higher computer equipped with an EGA or VGA display adapter and MS-DOS 3.00 or above operational system. The chromatogram printout can be performed on any Epson-compatible dot matrix printer. In this work, Epson FX-1050 and FX-800 nine-pin matrix printers were used to test the printing facilities. The software requires at least 300 K of RAM and about 200 K of free disk space; it can be installed on a hard disk or a floppy disk. A co-processor is preferable, but the software can emulate its work to accelerate floating point operations for the systems without a mathematical co-processor. The graphical image of a chromatogram can be stored on the disk or loaded for the future use with "Chromatogram Generator" or with other software programs. The size of the stored file is minimal (not more than 300 bytes), which allows the storage of large amount of data on the disk (approximately 2000 files per double-sided high-density floppy disk).

3. Peak shapes

In this software, we applied the approach of chromatogram creation from a certain set of various predefined function templates. The user can visually determine the parameters of each peak, and the chromatogram is treated as the superposition of the user-defined functions. Theoretically, symmetrical, well resolved peaks can be described with sufficient accuracy by the approximation with a Gaussian distribution curve [6]:

$$h(t) = h_{max} \exp\left[-\frac{1}{2}\left(\frac{t - t_0}{\sigma}\right)^2 \right] \quad (1)$$

where h_{max} is the maximum height of the peak, t_0 is the time corresponding to the peak minimum and σ is the standard deviation.

However, usually real peaks are asymmetric owing to non-linear sorption isotherms or other causes. Hence it is convenient to treat each peak as the combination of two separate (front and tail) parts and to consider two separate equations for each part:

$$h(t) = \begin{cases} \dfrac{h_{max}}{\sigma_1\sqrt{2\pi}} \cdot \exp\left[-\dfrac{(t - t_0)^2}{2\sigma_1^2} \right] & -\infty < t \leq t_0 \\ \dfrac{h_{max}}{\sigma_2\sqrt{2\pi}} \cdot \exp\left[-\dfrac{(t - t_0)^2}{2\sigma_2^2} \right] & t_0 \leq t < +\infty \end{cases} \quad (2)$$

In the above equations the peak is described by the curve of a Gaussian distribution with half-widths (σ_1 and σ_2) defined independently for the front and tail parts. This distribution is called bi-Gaussian [7]. Peaks with a strongly diffuse front are poorly approximated with Eq. 2, and a more accurate description is reached with the approximation

$$h(t) = \frac{h_{max}}{\tau} \int_{-\infty}^{t} dt_1 \cdot \exp\left[-\frac{(t_1 - t_0)^2}{2\omega^2} \right] \cdot \exp\left[-\frac{(t - t_1 + t_0)}{\tau} \right] \quad (3)$$

A tailed peak can also be described by an exponentially modified Gaussian function [8]:

$$h(t) = \frac{h_{max}}{\tau} \cdot \exp\left[\frac{1}{2}\left(\frac{\tau}{\sigma}\right)^2 - \frac{(t - t_0)}{\tau} \right] \cdot \int_{-\infty}^{z} \exp\left[\frac{\left(-\frac{x^2}{2}\right) dx}{\sqrt{2\pi}} \right]$$

$$z = \frac{1}{\sqrt{2}} \cdot \left(\frac{t}{\sigma} - \frac{\sigma}{\tau} \right) \quad (4)$$

where $x = (t - t_0)/\sigma$, h_{max} is the maximum height of the peak, t_0 is the time corresponding to the peak maximum and τ and σ are the parameters corresponding to the curvature and the width of the peak. In this case, varying the ratio of τ to σ can result in peaks with predefined overlapping (S) and resolution (R_S). This problem was described in detail by Sekulic and Haddad [9] and Crubner [10,11].

In addition to modified Gaussian curves, the chromatographic peaks can be described with a first-order Poisson distribution function (Eq. 5) and a Cauchi function (Eq. 6) [7,12]:

$$h(t) = h_{max} \cdot \exp[\lambda(e^{it} - 1)] \quad \lambda > 0 \tag{5}$$

$$h(t) = \frac{h_{max}}{1 + \left[\frac{2(t - t_0)}{\sigma}\right]^2} \tag{6}$$

In some cases (e.g., gas chromatographic determination of alcohols), the peaks can be treated with good accuracy as triangles. Strenberg [13] proposed the following equation to describe these peaks:

$$h(t) = \begin{cases} h_{max} \cdot \left[1 - \frac{|t - t_0|}{\sigma}\right] & |t - t_0| < \sigma \\ 0 & |t - t_0| \geq \sigma \end{cases} \tag{7}$$

This approach does not always provide the required accuracy of the approximation, but it is suitable for the purpose of the software discussed. The parabolic shape of the peaks is, of course, very peculiar in the "pure form", but it may be useful for modelling shouldered peaks.

4. Chromatogram modelling process

The program interface is shown in Fig. 1. To model the real chromatogram the user defines the peak shape, half-widths (separately for the left and right parts), peak height and retention time using the icons at the top of the screen. The ranges over which these parameters vary are presented in Table 1. This determines the "skeleton" of the chromatogram.

The user has an option to fine tune further to generate a chromatographic peak by choosing and editing any peak to obtain maximum agree-

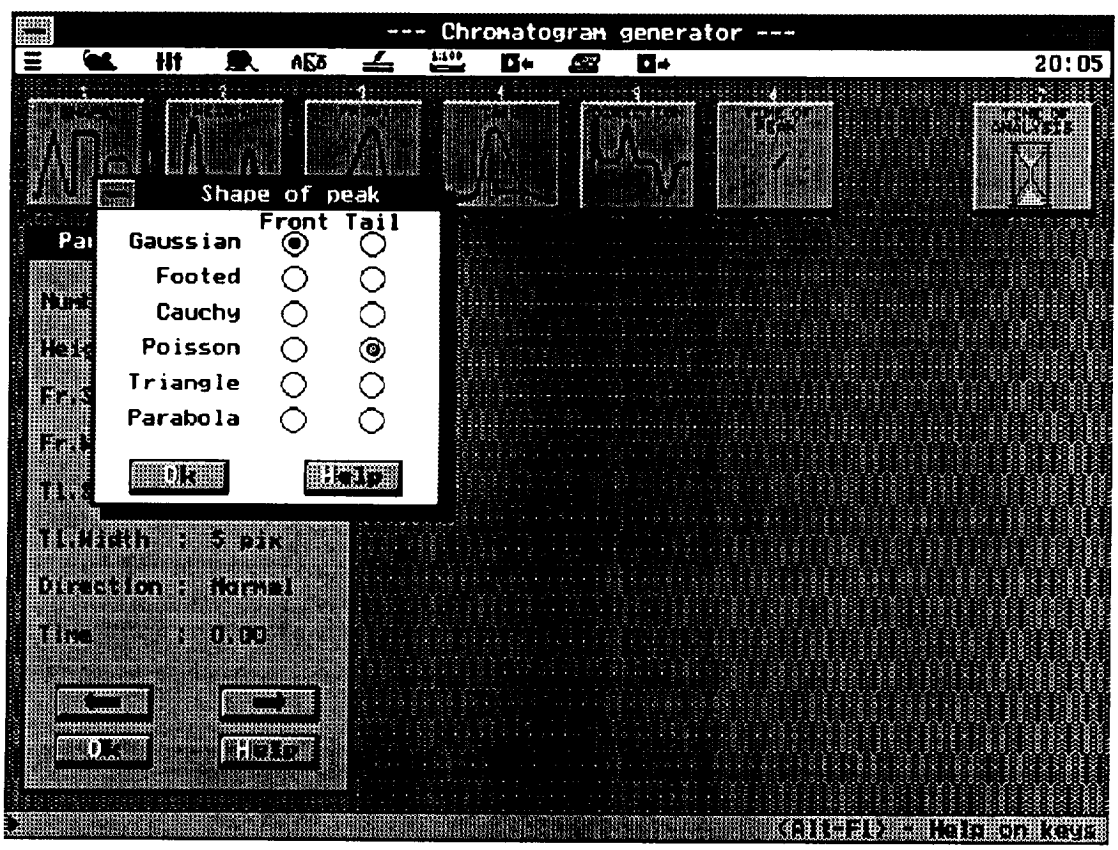

Fig. 1. The main window of the "Chromatogram Generator" program.

Table 1
Variables in the "Chromatogram Generator" software used to approximate chromatographic peaks

Varying parameter	Range of variation
Time of analysis	From 2 to 120 min
Retention time of peaks	With precision of 0.01 min
Height of peaks[a]	From 0 to 400 mm
Width of front or tail of peaks	From 0 to 200 mm
Direction of peaks	"Positive" or "negative"
Shape of peaks[b]	Six functions

[a] Height and width of peak can be defined in screen pixels also.
[b] Front and tail of peak can be approximated with different functions.

ment between the model and the real chromatograms (Fig. 2). The baseline drift can be emulated in most instances by the superposition of "pseudo-peaks", e.g., with very large widths. Usually, the process of modelling of a real chromatogram takes 2–3 min.

It was found that "Chromatogram Generator" provided a good approximation of real chromatograms (Figs. 3 and 4). The process of the construction of a series of model chromatograms for expert estimation is simplified by using a template stored on the disk. Users have the option to edit the files with chromatogram images using a simple additional utility program. This allows subsequent variation of a single parameter (e.g., peak width) to produce a series of chromatograms with various peak resolutions for subsequent expert estimation.

The approximation of chromatograms with the use of a number of functions leads to a considerable increase in the compression ratio of the

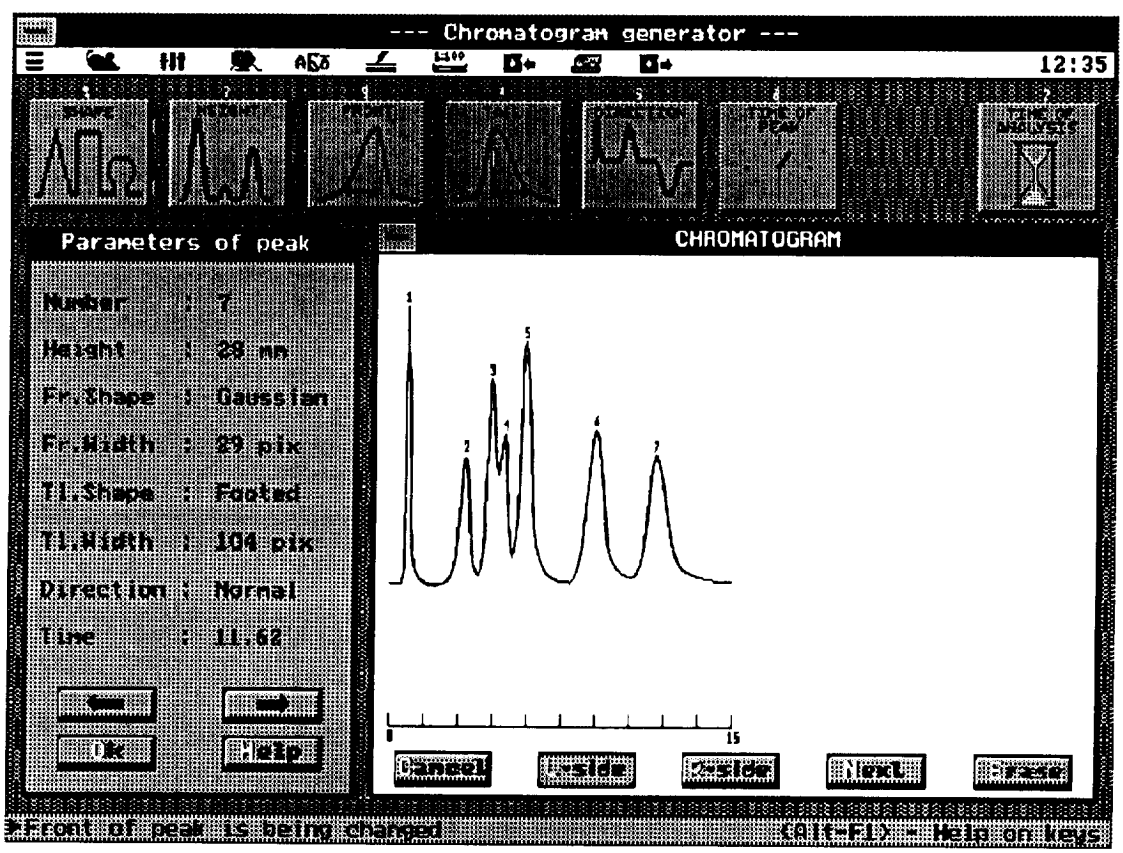

Fig. 2. Process of modelling real chromatograms.

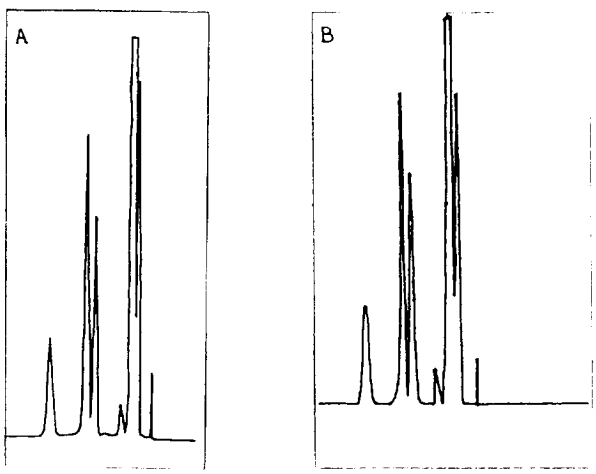

Fig. 3. (A) Real and (B) modelled chromatograms.

information. This decrease in the file size results from the approximation of chromatographic peaks by superposition of mathematical functions. Thus, "pixel-by-pixel" storage of a chromatographic image is not necessary. The size of the file of a chromatogram is 100–300 bytes. It can be added to a database, which allows one to keep the graphical image of the chromatogram. The file sizes for graphical images of chromatograms produced with some popular image-editing software packages and "Chromatogram Generator" are presented in Table 2. It should be noted that the file size generated by "Chromatogram Generator" is only about 168 bytes.

Table 2
File sizes with image of the chromatograms (Fig. 2) built by different graphic image editors and the "Chromatogram Generator" program

Graphic image editor	Size of file (bytes)
Ventura Publisher V.2.0	48 952
MS-Paintbrush V.4.0	42 537
Microsoft Word V.2.0c	23 582
Pizza V.1.12	5376
Chromatogram Generator	168

5. Conclusions

The developed software allows the solution of the following tasks: (i) approximation of chromatographic peaks with several mathematical functions and their superposition, which is of interest for several chromatographic methods; (ii) construction of series of chromatograms with various parameters (or with a smooth variation of a single or several parameters) to obtain a

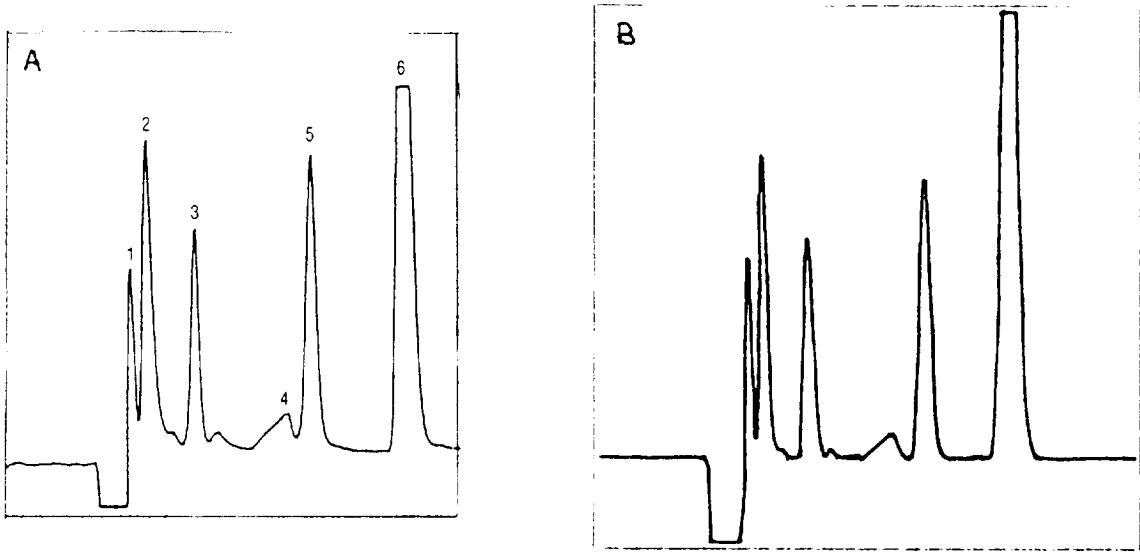

Fig. 4. (A) Real and (B) modelled chromatograms with different shapes and direction of peaks.

representative set for expert estimation of chromatogram quality, which is necessary for the choice of the weight factors of the equations of optimization criteria; and (iii) the construction of a graphical image of a chromatogram with minimal size and the ability to be added to a chromatographic database.

References

[1] M.W. Watson and P.W. Carr, *Anal. Chem.*, 51 (1979) 1835.
[2] J.L. Glajch, J.J. Kirkland, K.M. Squire and J.M. Minor, *J. Chromatogr.*, 199 (1980) 57.
[3] A.C.J.H. Drouen, H.A.H. Billiet, P.J. Schoenmakers and L. Galan, *Chromatographia*, 16 (1982) 48.
[4] J.C. Berridge and E.G. Morrisey, *J. Chromatogr.*, 316 (1984) 69.
[5] O.N. Obrezkov, A.V. Pirogov, I.V. Pletnev and O.A. Shpigun, *Mikrochim. Acta*, 1 (1991) 293.
[6] W.E. Barber and P.W. Carr, *Anal. Chem.*, 53 (1981) 1939.
[7] T.S. Buyeand and K. De Clerk, *Anal. Chem.*, 44 (1972) 1273.
[8] J.P. Foley and J.G. Dorsey, *Anal. Chem.*, 55 (1983) 730.
[9] S. Sekulic and P.R. Haddad, *J. Chromatogr.*, 459 (1988) 65.
[10] O. Crubner, *Adv. Chromatogr.*, 6 (1968) 173.
[11] O. Crubner, *Anal. Chem.*, 43 (1971) 1934.
[12] R.D.B. Fraser and E. Suzuki, *Anal. Chem.*, 38 (1966) 1770.
[13] J.C. Strenberg, *Adv. Chromatogr.*, 2 (1966) 205.